

1  
SEQUENCE LISTING

GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Max-Planck-Gesellschaft zur Förderung der  
Wissenschaften e.V.  
(B) STREET: none  
(C) CITY: Berlin  
(D) STATE: none  
(E) COUNTRY: Germany  
(F) POSTAL CODE (ZIP): none

- (A) NAME: National Public Health Institute  
(B) STREET: Mannerheimintie 166  
(C) CITY: Helsinki  
(D) STATE: none  
(E) COUNTRY: Finland  
(F) POSTAL CODE (ZIP): 00300

(ii) TITLE OF INVENTION: Novel nucleic acid molecule encoding a  
(poly)peptide co-segregating in mutated form with  
Autoimmune Polyendocrinopathy Candidiasis Ectodermal  
Dystrophy (APECED)

(iii) NUMBER OF SEQUENCES: 32

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2245 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CGGGCGCACA GCCGGCGCGG AGGCCCCACA GCCCCGCCGG GACCCGAGGC CAAGCGAGGG 60  
GCTGCCAGTG TCCCGGGACC CACCGCGTCC GCCCCAGCCC CGGGTCCCCG CGCCACCCCC 120  
ATGGCGACGG ACGCGGCGCT ACGCCGGCTT CTGAGGCTGC ACCGCACGGA GATCGGGGTG 180  
GCCGTGGACA GCGCCTTCCC ACTGCTGCAC GCGCTGGCTG ACCACGACGT GGTCCCCGAG 240  
GACAAGTTTC AGGAGACGCT TCATCTGAAG GAAAAGGAGG GCTGCCCCCA GGCCTTCCAC 300  
GCCCTCCTGT CCTGGCTGCT GACCCAGGAC TCCACAGCCA TCCTGGACTT CTGGAGGGTG 360

CTGTTCAAGG ACTACAACCT GGAGCGCTAT GGCCGGCTGC AGCCCATCCT GGACAGCTTC 420  
 CCCAAAGATG TGGACCTCAG CCAGCCCCGG AAGGGGAGGA AGCCCCCGGC CGTCCCCAAG 480  
 GCTTTGGTAC CGCCACCCAG ACTCCCCACC AAGAGGAAGG CCTCAGAAGA GGCTCGAGCT 540  
 GCCGCGCCAG CAGCCCTGAC TCCAAGGGGC ACCGCCAGCC CAGGCTCTCA ACTGAAGGCC 600  
 AAGCCCCCA AGAAGCCGGA GAGCAGCGCA GAGCAGCAGC GCCTTCCACT CGGGAACGGG 660  
 ATTCAGACCA TGTCAGCTTC AGTCCAGAGA GCTGTGGCCA TGTCTCCGG GGACGTCCCG 720  
 GGAGCCCGAG GGGCCGTGGA GGGGATCCTC ATCCAGCAGG TGTTTGAGTC AGGCGGCTCC 780  
 AAGAAGTGCA TCCAGGTTGG TGGGGAGTTC TACTCTCCA GCAAGTTCGA AGACTCCGGC 840  
 AGTGGGAAGA ACAAGGCCCG CAGCAGCAGT GGCCCGAAGC CTCTGGTTTC AGCCAAGGGA 900  
 GCCCAGGGCG CTGCCCCCGG TGGAGGTGAG GCTAGGCTGG GCCAGCAGGG CAGCGTTCCC 960  
 GCCCCTCTGG CCCTCCCCAG TGACCCCCAG CTCCACCAGA AGAATGAGGA CGAGTGTGCC 1020  
 GTGTGTCGGG ACGGCGGGGA GCTCATCTGC TGTGACGGCT GCCCTCGGGC CTTCCACCTG 1080  
 GCCTGCCTGT CCCCTCCGCT CCGGGAGATC CCCAGTGGGA CCTGGAGGTG CTCCAGCTGC 1140  
 CTGCAGGCAA CAGTCCAGGA GGTGCAGCCC CGGGCAGAGG AGCCCCGGCC CCAGGAGCCA 1200  
 CCCGTGGAGA CCCCGCTCCC CCCGGGGCTT AGGTCGGCGG GAGAGGAGGT AAGAGGTCCA 1260  
 CCTGGGGAAC CCCTAGCCGG CATGGACACG ACTCTTGTCT ACAAGCACCT GCCGGCTCCG 1320  
 CCTTCTGCAG CCCCGCTGCC AGGGCTGGAC TCCTCGGCCC TGCACCCCCT ACTGTGTGTG 1380  
 GGTCTGAGG GTCAGCAGAA CCTGGCTCCT GGTGCGCGTT GCGGGGTGTG CGGAGATGGT 1440  
 ACGGACGTGC TCGGCTGTAC TCACTGCGCC GCTGCCTTTC ACTGGCGCTG CCACTTCCCA 1500  
 GCCGGCACCT CCCGGCCCCG GACGGGCCTG CGCTGCAGAT CCGTCTCAGG AGACGTGACC 1560  
 CCAGCCCCTG TGGAGGGGGT GCTGGCCCCC AGCCCCGCCC GCCTGGCCCC TGGGCTGCC 1620  
 AAGGATGACA CTGCCAGTCA CGAGCCCGCT CTGCACAGGG ATGACCTGGA GTCCCTTCTG 1680  
 AGCGAGCACA CCTTCGATGG CATCCTGCAG TGGGCCATCC AGAGCATGGC CCGTCCGGCG 1740  
 GCCCCCTTCC CCTCCTGACC CCAGATGGCC GGGACATGCA GCTCTGATGA GAGAGTGCTG 1800  
 AGAAGGACAC CTCCTTCCTC AGTCCTGGAA GCCGGCCGGC TGGGATCAAG AAGGGGACAG 1860  
 CGCCACCTCT TGTCAGTGCT CGGCTGTAAA CAGCTCTGTG TTTCTGGGGA CACCAGCCAT 1920  
 CATGTGCCTG GAAATTAAAC CCTGCCCCAC TTCTCTACTC TGAAGTCCC CGGAGCCTC 1980  
 TCCTTGCTG GTGACCTACT AAAAATATAA AAATTAGCTG GGTGTGGTGG TGGGTGCCTG 2040  
 TAATCCCAGC TACATGGGAG CCTGAGGCAT GAGAATCACT TGAACCTGGG AGGTGGAGGT 2100  
 TGCAGTGAGC TGAGATTGCG CCACTGCACT CCAGTCTGGT CGGCAAGAGT GAGACTCCGT 2160  
 CTCAAAAACA AAACAAAAAA ACCACATAAC ATAAATTTAT CATCTCGACC ACTTTTCAGT 2220

B/  
 nt

TCAGTGGCAT TCACATCTCA TGTA

2245

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 545 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg Thr  
 1 5 10 15  
 Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu Leu His Ala Leu  
 20 25 30  
 Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu His  
 35 40 45  
 Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu Ser  
 50 55 60  
 Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp Phe Trp Arg Val  
 65 70 75 80  
 Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg Leu Gln Pro Ile  
 85 90 95  
 Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln Pro Arg Lys Gly  
 100 105 110  
 Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro Pro Pro Arg Leu  
 115 120 125  
 Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala Ala Ala Pro Ala  
 130 135 140  
 Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser Gln Leu Lys Ala  
 145 150 155 160  
 Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln Gln Arg Leu Pro  
 165 170 175  
 Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val Gln Arg Ala Val  
 180 185 190  
 Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly Ala Val Glu Gly  
 195 200 205  
 Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser Lys Lys Cys Ile  
 210 215 220  
 Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe Glu Asp Ser Gly

225                      230                      235                      240  
 Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro Lys Pro Leu Val  
                                  245                      250                      255  
  
 Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly Gly Glu Ala Arg  
                                  260                      265                      270  
  
 Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala Leu Pro Ser Asp  
                                  275                      280                      285  
  
 Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp  
                                  290                      295                      300  
  
 Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu  
 305                                   310                      315                      320  
  
 Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg  
                                  325                      330                      335  
  
 Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala  
                                  340                      345                      350  
  
 Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro  
                                  355                      360                      365  
  
 Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro  
                                  370                      375                      380  
  
 Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro  
 385                                   390                      395                      400  
  
 Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro  
                                  405                      410                      415  
  
 Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala  
                                  420                      425                      430  
  
 Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His  
                                  435                      440                      445  
  
 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Ala Gly Thr Ser  
                                  450                      455                      460  
  
 Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr  
 465                                   470                      475                      480  
  
 Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala  
                                  485                      490                      495  
  
 Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu Pro Ala Leu His  
                                  500                      505                      510  
  
 Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile  
                                  515                      520                      525  
  
 Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro  
                                  530                      535                      540  
  
 Ser

B1  
 nk

545

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ACAGGCAGGC AG

12

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ACAGGCAGGC CA

12

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCAGGCCAGG TG

12

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGGAGGGGA CA

12

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTGTGGACTG TCACGGAAAC CCCACGTGT GATGGAAAGT CCAAATTCT ACAGGAGTCT

60

TTCTGTTGAT CTCCAGTCAG AGGCTGGGGG

90

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAGGGGCTGG TGTGGAAAGC CCCACGGCAT GGTGGAAAGT CCGAAATTCT ACAGGGGCCT

60

CTTTGTAAA CCTCCATGCA AGAGGCTGGG

90

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

NNGNGGNNG TNNNGNAANC CCCNNNGNNT GNTGGAAAGT CCNAAATTCT ACAGGNGNCT 60  
 NTNTGTTNAN CNNC NNTNNN AGNNNNNGGG 90

N represents any nucleotide or a gap

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1659 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ATGGCAGGTG GGGATGGAAT GCTACGCCGT CTGCTGAGGC TGCACCGCAC CGAGATCGCG 60  
 GTGGCCATAG ACAGTGCCTT TCCGCTGCTG CATGCTCTAG CCGACCACGA CGTGGTCCCT 120  
 GAGGACAAGT TCCAGGAGAC GCTCCGTCTG AAGGAGAAGG AAGGCTGCCC CCAGGCCTTC 180  
 CACGCCCTGC TGTCTGGCT CCTGACCCGG GACAGTGGGG CCATCCTGGA TTTCTGGAGG 240  
 ATTCTCTTTA AGGACTACAA TCTGGAGCGG TACAGCCGCC TGCATAGCAT CCTGGACGGC 300  
 TTCCCAAAG ATGTGGACCT AAACCAGTCC CGGAAAGGGA GAAAGCCCCT TGCTGGTCCC 360  
 AAGGCCGCGG TACTGCCACC CAGACCCCCC ACCAAGAGAA AAGCACTGGA GGAGCCTCGA 420  
 GCCACCCAC CAGCAACTCT GGCTCAAAG AGCGTCTCCA GCCCAGGCTC CCACCTGAAG 480  
 ACTAAGCCCC CTAAGAAGCC AGATGGCAAC TTGGAGTCAC AGCACCTTCC TCTTGAAAC 540  
 GGAATTCAGA CCATGGCAGC TTCTGTCCAG AGAGCTGTGA CCGTGGCCTC TGGGGATGTT 600  
 CCAGGAACCC GAGGGGCCGT GGAAGGGATC CTTATCCAGC AGGTGTTTGA GTCAGGAAGA 660  
 TCCAAGAAGT GCATTCAGGT TGGGGGAGAG TTTTATACAC CCAACAAGTT CGAAGACCCC 720  
 AGTGGCAATT TGAAGAACAA GGCCCGAGT GGTAGCAGCC TAAAGCCAGT GGTCCGAGCC 780  
 AAGGGAGCCC AGGTCACTAT ACCTGGTAGA GATGAGCAGA AAGTGGGCCA GCAGTGTGGG 840  
 GTTCCTCCCC TTCCATCCCT CCCAGTGAG CCCAGGTTA ACCAGAAGAA CGAGGATGAG 900  
 TGTGCCGTGT GCCACGACGG AGGTGAGCTC ATCTGTTGTG ACGGCTGTCC CCGGGCCTTC 960  
 CACCTGGCTT GCCTGTCCCC ACCTCTGCAG GAGATCCCCA GTGGCCTCTG GAGATGCTCC 1020  
 TGCTGCCTCC AGGGCAGAGT CCAACAGAAC CTGTCCCAGC CTGAGGTGTC CAGGCCCCCG 1080  
 GAGCTACCTG CAGAGACCCC GATCCTCGTG GGACTGAGGT CAGCTTCAGA GAAAACCAGG 1140

GGCCCATCCA GGGAGCTCAA AGCCAGCTCT GATGCTGCTG TCACATATGT GAACCTGCTG 1200  
GCCCCGCACC CTGCAGCTCC TCTGCTGGAG CCTTCAGCAC TGTGCCCTCT ACTGAGTGCT 1260  
GGGAATGAGG GGC GGCCAGG TCCAGCACCA AGCGCGCGAT GCAGTGTGTG TGGCGATGGC 1320  
ACCGAGGTGT TGC GGTGTGC ACAC TGTGCC GCTGCCTTCC ACTGGCGCTG CCACTTCCCC 1380  
ACGGCCGCCG CCCGGCCGGG GACCAATCTC CGCTGCAAAT CCTGCTCTGC AGACTCGACT 1440  
CCCACGCCAG GCACACCGGG CGAAGCTGTA CCCACCTCTG GGCCCCGTCC AGCACCTGGG 1500  
CTTGCCAAGG TAGGGGACGA CTCTGCTAGT CACGACCCTG TTCTACATAG GGACGACCTG 1560  
GAGTCCCTCC TCAATGAGCA CTCATTGAC GGCATCCTGC AGTGGGCCAT CCAGAGCATG 1620  
TCACGCCCCG TGGCCGAGAC ACCACCCTTC TCTTCCTGA 1659

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 552 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Ala Gly Gly Asp Gly Met Leu Arg Arg Leu Leu Arg Leu His Arg  
1 5 10 15  
Thr Glu Ile Ala Val Ala Ile Asp Ser Ala Phe Pro Leu Leu His Ala  
20 25 30  
Leu Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu  
35 40 45  
Arg Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu  
50 55 60  
Ser Trp Leu Leu Thr Arg Asp Ser Gly Ala Ile Leu Asp Phe Trp Arg  
65 70 75 80  
Ile Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Ser Arg Leu His Ser  
85 90 95  
Ile Leu Asp Gly Phe Pro Lys Asp Val Asp Leu Asn Gln Ser Arg Lys  
100 105 110  
Gly Arg Lys Pro Leu Ala Gly Pro Lys Ala Ala Val Leu Pro Pro Arg  
115 120 125  
Pro Pro Thr Lys Arg Lys Ala Leu Glu Glu Pro Arg Ala Thr Pro Pro  
130 135 140  
Ala Thr Leu Ala Ser Lys Ser Val Ser Ser Pro Gly Ser His Leu Lys  
145 150 155 160



Thr Lys Pro Pro Lys Lys Pro Asp Gly Asn Leu Glu Ser Gln His Leu  
 165 170 175  
 Pro Leu Gly Asn Gly Ile Gln Thr Met Ala Ala Ser Val Gln Arg Ala  
 180 185 190  
 Val Thr Val Ala Ser Gly Asp Val Pro Gly Thr Arg Gly Ala Val Glu  
 195 200 205  
 Gly Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Arg Ser Lys Lys Cys  
 210 215 220  
 Ile Gln Val Gly Gly Glu Phe Tyr Thr Pro Asn Lys Phe Glu Asp Pro  
 225 230 235 240  
 Ser Gly Asn Leu Lys Asn Lys Ala Arg Ser Gly Ser Ser Leu Lys Pro  
 245 250 255  
 Val Val Arg Ala Lys Gly Ala Gln Val Thr Ile Pro Gly Arg Asp Glu  
 260 265 270  
 Gln Lys Val Gly Gln Gln Cys Gly Val Pro Pro Leu Pro Ser Leu Pro  
 275 280 285  
 Ser Glu Pro Gln Val Asn Gln Lys Asn Glu Asp Glu Cys Ala Val Cys  
 290 295 300  
 His Asp Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe  
 305 310 315 320  
 His Leu Ala Cys Leu Ser Pro Pro Leu Gln Glu Ile Pro Ser Gly Leu  
 325 330 335  
 Trp Arg Cys Ser Cys Cys Leu Gln Gly Arg Val Gln Gln Asn Leu Ser  
 340 345 350  
 Gln Pro Glu Val Ser Arg Pro Pro Glu Leu Pro Ala Glu Thr Pro Ile  
 355 360 365  
 Leu Val Gly Leu Arg Ser Ala Ser Glu Lys Thr Arg Gly Pro Ser Arg  
 370 375 380  
 Glu Leu Lys Ala Ser Ser Asp Ala Ala Val Thr Tyr Val Asn Leu Leu  
 385 390 395 400  
 Ala Pro His Pro Ala Ala Pro Leu Leu Glu Pro Ser Ala Leu Cys Pro  
 405 410 415  
 Leu Leu Ser Ala Gly Asn Glu Gly Arg Pro Gly Pro Ala Pro Ser Ala  
 420 425 430  
 Arg Cys Ser Val Cys Gly Asp Gly Thr Glu Val Leu Arg Cys Ala His  
 435 440 445  
 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Thr Ala Ala Ala  
 450 455 460  
 Arg Pro Gly Thr Asn Leu Arg Cys Lys Ser Cys Ser Ala Asp Ser Thr  
 465 470 475 480

B1

Pro Thr Pro Gly Thr Pro Gly Glu Ala Val Pro Thr Ser Gly Pro Arg  
485 490 495

Pro Ala Pro Gly Leu Ala Lys Val Gly Asp Asp Ser Ala Ser His Asp  
500 505 510

Pro Val Leu His Arg Asp Asp Leu Glu Ser Leu Leu Asn Glu His Ser  
515 520 525

Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ser Arg Pro Leu  
530 535 540

Ala Glu Thr Pro Pro Phe Ser Ser  
545 550

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 545 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg Thr  
1 5 10 15

Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu Leu His Ala Leu  
20 25 30

Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu His  
35 40 45

Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu Ser  
50 55 60

Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp Phe Trp Arg Val  
65 70 75 80

Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg Leu Gln Pro Ile  
85 90 95

Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln Pro Arg Lys Gly  
100 105 110

Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro Pro Pro Arg Leu  
115 120 125

Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala Ala Ala Pro Ala  
130 135 140

Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser Gln Leu Lys Ala  
145 150 155 160

Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln Gln Arg Leu Pro  
165 170 175

Bf

Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val Gln Arg Ala Val  
 180 185 190  
 Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly Ala Val Glu Gly  
 195 200 205  
 Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser Lys Lys Cys Ile  
 210 215 220  
 Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe Glu Asp Ser Gly  
 225 230 235 240  
 Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro Lys Pro Leu Val  
 245 250 255  
 Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly Gly Glu Ala Arg  
 260 265 270  
 Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala Leu Pro Ser Asp  
 275 280 285  
 Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp  
 290 295 300  
 Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu  
 305 310 315 320  
 Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg  
 325 330 335  
 Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala  
 340 345 350  
 Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro  
 355 360 365  
 Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro  
 370 375 380  
 Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro  
 385 390 395 400  
 Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro  
 405 410 415  
 Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala  
 420 425 430  
 Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His  
 435 440 445  
 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Ala Gly Thr Ser  
 450 455 460  
 Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr  
 465 470 475 480  
 Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala  
 485 490 495

B/L

Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu Pro Ala Leu His  
 500 505 510

Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile  
 515 520 525

Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro  
 530 535 540

Ser  
 545

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 552 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Met Ala Gly Gly Asp Gly Met Leu Arg Arg Leu Leu Arg Leu His Arg  
 1 5 10 15

Thr Glu Ile Ala Val Ala Ile Asp Ser Ala Phe Pro Leu Leu His Ala  
 20 25 30

Leu Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu  
 35 40 45

Arg Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu  
 50 55 60

Ser Trp Leu Leu Thr Arg Asp Ser Gly Ala Ile Leu Asp Phe Trp Arg  
 65 70 75 80

Ile Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Ser Arg Leu His Ser  
 85 90 95

Ile Leu Asp Gly Phe Pro Lys Asp Val Asp Leu Asn Gln Ser Arg Lys  
 100 105 110

Gly Arg Lys Pro Leu Ala Gly Pro Lys Ala Ala Val Leu Pro Pro Arg  
 115 120 125

Pro Pro Thr Lys Arg Lys Ala Leu Glu Glu Pro Arg Ala Thr Pro Pro  
 130 135 140

Ala Thr Leu Ala Ser Lys Ser Val Ser Ser Pro Gly Ser His Leu Lys  
 145 150 155 160

Thr Lys Pro Pro Lys Lys Pro Asp Gly Asn Leu Glu Ser Gln His Leu  
 165 170 175

Pro Leu Gly Asn Gly Ile Gln Thr Met Ala Ala Ser Val Gln Arg Ala  
 180 185 190  
 Val Thr Val Ala Ser Gly Asp Val Pro Gly Thr Arg Gly Ala Val Glu  
 195 200 205  
 Gly Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Arg Ser Lys Lys Cys  
 210 215 220  
 Ile Gln Val Gly Gly Glu Phe Tyr Thr Pro Asn Lys Phe Glu Asp Pro  
 225 230 235 240  
 Ser Gly Asn Leu Lys Asn Lys Ala Arg Ser Gly Ser Ser Leu Lys Pro  
 245 250 255  
 Val Val Arg Ala Lys Gly Ala Gln Val Thr Ile Pro Gly Arg Asp Glu  
 260 265 270  
 Gln Lys Val Gly Gln Gln Cys Gly Val Pro Pro Leu Pro Ser Leu Pro  
 275 280 285  
 Ser Glu Pro Gln Val Asn Gln Lys Asn Glu Asp Glu Cys Ala Val Cys  
 290 295 300  
 His Asp Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe  
 305 310 315 320  
 His Leu Ala Cys Leu Ser Pro Pro Leu Gln Glu Ile Pro Ser Gly Leu  
 325 330 335  
 Trp Arg Cys Ser Cys Cys Leu Gln Gly Arg Val Gln Gln Asn Leu Ser  
 340 345 350  
 Gln Pro Glu Val Ser Arg Pro Pro Glu Leu Pro Ala Glu Thr Pro Ile  
 355 360 365  
 Leu Val Gly Leu Arg Ser Ala Ser Glu Lys Thr Arg Gly Pro Ser Arg  
 370 375 380  
 Glu Leu Lys Ala Ser Ser Asp Ala Ala Val Thr Tyr Val Asn Leu Leu  
 385 390 395 400  
 Ala Pro His Pro Ala Ala Pro Leu Leu Glu Pro Ser Ala Leu Cys Pro  
 405 410 415  
 Leu Leu Ser Ala Gly Asn Glu Gly Arg Pro Gly Pro Ala Pro Ser Ala  
 420 425 430  
 Arg Cys Ser Val Cys Gly Asp Gly Thr Glu Val Leu Arg Cys Ala His  
 435 440 445  
 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Thr Ala Ala Ala  
 450 455 460  
 Arg Pro Gly Thr Asn Leu Arg Cys Lys Ser Cys Ser Ala Asp Ser Thr  
 465 470 475 480  
 Pro Thr Pro Gly Thr Pro Gly Glu Ala Val Pro Thr Ser Gly Pro Arg  
 485 490 495

Pro Ala Pro Gly Leu Ala Lys Val Gly Asp Asp Ser Ala Ser His Asp  
                   500                                  505                                  510

Pro Val Leu His Arg Asp Asp Leu Glu Ser Leu Leu Asn Glu His Ser  
                   515                                  520                                  525

Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ser Arg Pro Leu  
                   530                                  535                                  540

Ala Glu Thr Pro Pro Phe Ser Ser  
                   545                                  550

## (2) INFORMATION FOR SEQ ID NO: 14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 550 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Xaa Xaa Xaa Asp Xaa Xaa Leu Arg Arg Leu Leu Arg Leu His Arg Thr  
   1                  5                                  10                                  15

Glu Ile Ala Val Ala Xaa Asp Ser Ala Phe Pro Leu Leu His Ala Leu  
                   20                                  25                                  30

Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu Xaa  
                   35                                  40                                  45

Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu Ser  
                   50                                  55                                  60

Trp Leu Leu Thr Xaa Asp Ser Xaa Ala Ile Leu Asp Phe Trp Arg Xaa  
                   65                                  70                                  75                                  80

Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Xaa Arg Leu Xaa Xaa Ile  
                   85                                  90                                  95

Leu Asp Xaa Phe Pro Lys Asp Val Asp Leu Xaa Gln Xaa Arg Lys Gly  
                   100                                  105                                  110

Arg Lys Pro Xaa Ala Xaa Pro Lys Ala Xaa Val Xaa Pro Pro Arg Xaa  
                   115                                  120                                  125

Pro Thr Lys Arg Lys Ala Xaa Glu Glu Xaa Arg Ala Xaa Xaa Pro Ala  
                   130                                  135                                  140

Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Ser Pro Gly Ser Xaa Leu Lys Xaa  
                   145                                  150                                  155                                  160

Lys Pro Pro Lys Lys Pro Xaa Xaa Xaa Xaa Glu Xaa Gln Xaa Leu Pro  
                   165                                  170                                  175

Leu Gly Asn Gly Ile Gln Thr Met Xaa Ala Ser Val Gln Arg Ala Val  
 180 185 190  
 Xaa Xaa Xaa Ser Gly Asp Val Pro Gly Xaa Arg Gly Ala Val Glu Gly  
 195 200 205  
 Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Xaa Ser Lys Lys Cys Ile  
 210 215 220  
 Gln Val Gly Gly Glu Phe Tyr Thr Pro Xaa Lys Phe Glu Asp Xaa Ser  
 225 230 235 240  
 Gly Xaa Xaa Lys Asn Lys Ala Arg Ser Xaa Ser Xaa Xaa Lys Pro Xaa  
 245 250 255  
 Val Arg Ala Lys Gly Ala Gln Xaa Xaa Xaa Pro Gly Xaa Xaa Glu Xaa  
 260 265 270  
 Xaa Xaa Gly Gln Gln Xaa Xaa Val Pro Xaa Xaa Xaa Xaa Leu Pro Ser  
 275 280 285  
 Xaa Pro Gln Xaa Xaa Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Xaa  
 290 295 300  
 Asp Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His  
 305 310 315 320  
 Leu Ala Cys Leu Ser Pro Pro Leu Xaa Glu Ile Pro Ser Gly Xaa Trp  
 325 330 335  
 Arg Cys Ser Xaa Cys Leu Gln Xaa Xaa Val Gln Xaa Xaa Xaa Xaa Xaa  
 340 345 350  
 Xaa Glu Xaa Xaa Arg Pro Xaa Glu Xaa Pro Xaa Glu Thr Pro Xaa Xaa  
 355 360 365  
 Xaa Gly Leu Arg Ser Ala Xaa Glu Xaa Xaa Arg Gly Pro Xaa Xaa Glu  
 370 375 380  
 Xaa Xaa Ala Xaa Xaa Asp Xaa Xaa Xaa Xaa Tyr Xaa Xaa Leu Xaa Ala  
 385 390 395 400  
 Pro Xaa Xaa Ala Ala Pro Leu Xaa Xaa Leu Xaa Xaa Ser Ala Leu Xaa  
 405 410 415  
 Pro Leu Leu Xaa Xaa Gly Xaa Glu Gly Xaa Xaa Xaa Xaa Ala Pro Xaa  
 420 425 430  
 Ala Arg Cys Xaa Val Cys Gly Asp Gly Thr Xaa Val Leu Arg Cys Xaa  
 435 440 445  
 His Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Xaa Xaa Xaa  
 450 455 460  
 Xaa Arg Pro Gly Thr Xaa Leu Arg Cys Xaa Ser Cys Ser Xaa Asp Xaa  
 465 470 475 480  
 Thr Pro Xaa Pro Xaa Xaa Xaa Gly Xaa Xaa Xaa Pro Xaa Ser Xaa Xaa  
 485 490 495

Bl  
 wj

Arg Xaa Ala Pro Gly Xaa Ala Lys Xaa Xaa Asp Asp Xaa Ala Ser His  
 500 505 510

Xaa Pro Xaa Leu His Arg Asp Asp Leu Glu Ser Leu Leu Xaa Glu His  
 515 520 525

Xaa Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Xaa Arg Pro  
 530 535 540

Xaa Ala Xaa Xaa Pro Xaa  
 545 550

Xaa denotes any amino acid

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGGGCCTCGA TGGACGTCTC TGGGGCCCAG GTCGTGGTTC GCGCGCTA

48

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Pro Glu Leu Pro Ala Glu Thr Pro Gly Pro Ala Pro Ser Ala Arg  
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AGTGAGCCCC AGGTTAACCA GAACGAGGAT GAGTGTGCCG TGT

43

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Ser Glu Pro Gln Val Asn Gln Asn Glu Asp Glu Cys Ala Val  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GTCACCAGGC TCGGTTCCCT CGGGTCCCAT CTCTACTCGT CTTTCACC

48

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Val Val Arg Ala Lys Gly Ala Gln Gly Arg Asp Glu Gln Lys Val  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

AGAAGTGCAT CCAGGTTGGC

20

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GGAAGAGGGG CGTCAGCAAT

20

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 15 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg  
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 14 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Ser Gln Pro Arg Lys Gly Arg Lys Pro Pro Ala Val Pro Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Gly Gly Cys Thr Thr Cys Thr Gly Ala Gly Gly Cys Thr Gly Cys Ala  
1 5 10 15

Cys Cys

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GCTCTGGATG GCCTACTGC

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

AGAAGTGCAT CCAGGTTGGC

20

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GTGTGCTCGC TCAGAAGGG

19

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

TGGCAGGTGG GGATGGAA

18

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GGAGGGATGG AAGGGGAGGA

20

(2) INFORMATION FOR SEQ ID NO: 31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES


## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

TCCCACCTGA AGACTAAGC

19

## (2) INFORMATION FOR SEQ ID NO: 32:

## (i) SEQUENCE CHARACTERISTICS:

- 
- (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

## (iii) HYPOTHETICAL: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

TCACAGCTCT CTGGACAGAA

1

20

1